Выбор оптимальной модели в задаче моделирования динамики физической системы

Северилов Павел

Московский физико-технический институт
Кафедра интеллектуальных систем

Научный руководитель д.ф.-м.н. В. В. Стрижов

Москва,
2022 г.

Введение

Решается задача выбора оптимальной модели предсказания динамики физической системы. Под динамикой системы понимается изменение во времени параметров системы.

Проблема: Нейронные сети не имеют априорных знаний о моделируемой системе, что не позволяет получить оптимальные параметры, учитывающие физические законы.

Гипотеза: Внесение априорного знания о физике системы повышает качество модели.

Предлагается: модификация Лагранжевой нейронной сети (LNN), учитывающая трансляционную и вращательную симметрии кроме закона сохранения энергии. Проверить оптимальность LNN, сравнив с моделями без априорных знаний о моделируемой физической системе.

Эксперимент: Сгенерировать данные для системы двойного маятника и сравнить результаты моделирования динамики системы различными моделями нейронных сетей.

Постановка задачи регрессии динамики физической системы

1. Дана выборка m траекторий
$$\{\boldsymbol{x}_i, \boldsymbol{y}_i\}_{i=1}^{m},$$
где $\boldsymbol{x}_i = (\boldsymbol{q}_i, \dot{\boldsymbol{q}}_i)$ – координаты траектории движения двойного маятника, $\boldsymbol{y}_i = \dot{\boldsymbol{x}}_i = (\dot{\boldsymbol{q}}_i, \ddot{\boldsymbol{q}}_i)$ – динамика движения системы двойного маятника, $\boldsymbol{q}_i \in \mathbb{R}^{r\times n}$, где r – количество координат, n – длина траектории
2. Модель выбирается из класса нейронных сетей
$$\{\boldsymbol{f}_k\colon (\boldsymbol{w}, \boldsymbol{X}) \to \hat{\boldsymbol{y}} \mid k \in \mathcal{K}\},$$
где $\boldsymbol{w} \in \mathbb{W}$ – параметры модели,
$\hat{\boldsymbol{y}} = \boldsymbol{f}(\boldsymbol{X}, \boldsymbol{w}) \in \mathbb{R}^{2\times r\times n}, \boldsymbol{X} = \bigcup_{i=1}^{m} \boldsymbol{x}_i$.
3. Функция ошибки
$$\mathcal{L}(\boldsymbol{y}, \boldsymbol{X}, \boldsymbol{w}) = \|\hat{\boldsymbol{y}} - \boldsymbol{y}\|_2^2.$$
4. Решается задача оптимизации:
$$\mathbf{w}^* = \underset{\mathrm{w}\in\mathbb{W}}{\operatorname{argmin}}\big(\mathcal{L}(\mathbf{w})\big).$$

Система двойного маятника

1. Координаты $\mathbf{q} = (\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$ – углы между стержнями маятника и вертикальной осью;
2. **Лагранжиан** системы двойного маятника $L = T - V$, где T, V – кинетическая и потенциальная энергии системы:

$$L = \frac{1}{2}(m_1 + m_2) l_1^2 \dot{\theta}_1^2 + \frac{1}{2} m_2 l_2^2 \dot{\theta}_2^2 + m_2 l_1 l_2 \dot{\theta}_1 \dot{\theta}_2 \cos(\theta_1 - \theta_2) + (m_1 + m_2) g l_1 \cos\theta_1 + m_2 g l_2 \cos\theta_2$$

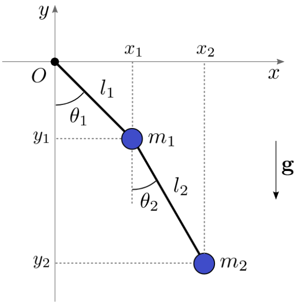


Схема физической системы двойного маятника

Лагранжева динамика

1. Лагранжиан системы L учитывает законы сохранения энергии, импульса и момента импульса;
2. Уравнение Эйлера Лагранжа описывает динамику системы
$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_j} = \frac{\partial L}{\partial q_j};$$
3. В нейронную сеть
$$f : \boldsymbol{X} = (\boldsymbol{q}, \dot{\boldsymbol{q}}) \to \boldsymbol{Y}$$
предлагается добавить априорные знания о физике системы, учитывая лагранжиан системы:
$$f : \boldsymbol{X} = (\boldsymbol{q}, \dot{\boldsymbol{q}}) \to \boldsymbol{L};$$
4. Из ограничений Эйлера-Лагранжа получаем выражение для обратного распространения ошибки[1]
$$\ddot{\boldsymbol{q}} = \left(\nabla_{\dot{\boldsymbol{q}}}\nabla_{\dot{\boldsymbol{q}}}^{\top} L\right)^{-1}\left[\nabla_{\boldsymbol{q}} L - \left(\nabla_{\boldsymbol{q}}\nabla_{\dot{\boldsymbol{q}}}^{\top} L\right)\dot{\boldsymbol{q}}\right], \quad \left(\nabla_{q}\nabla_{\dot{q}}^{\top} L\right)_{ij} = \frac{\partial^2 L}{\partial q_j \partial \dot{q}_i}$$

[1] *Cranmer, M. Greydanus et al.* Lagrangian Neural Networks. / ICLR 2020 Workshop

Лагранжевы нейронные сети (LNN)

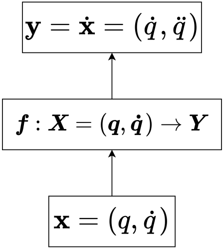


Схема работы базового решения нейронными сетями

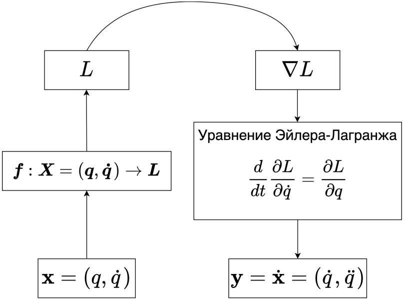


Схема работы LNN для моделирования динамики физической системы

LNN, учитывающая трансляционную и вращательную симметрии

Предлагается *Нётеровская Лагранжева нейронная сеть (LNN)*, которая получает на вход разницу между координатами

$$\Delta q_{ij} = \sqrt{(q_i - q_j) \cdot (q_i - q_j)}$$

и аппроксимирует потенциальную энергию системы $\boldsymbol{V}(\Delta \boldsymbol{q})$:

$$f : \boldsymbol{X} = (\boldsymbol{\Delta q}) \rightarrow \boldsymbol{V},$$

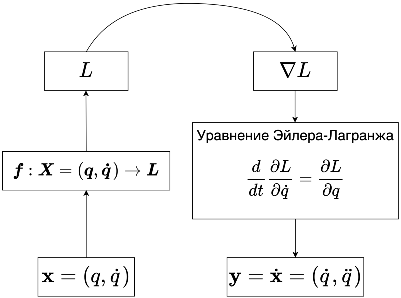


Схема работы LNN

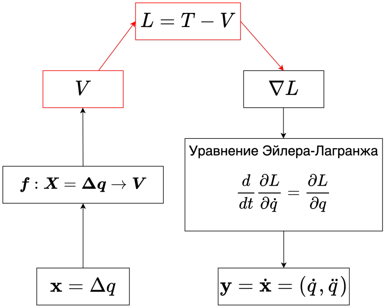


Схема работы Нётеровской LNN

Нётеровская Лагранжева нейронная сеть

Для системы двойного маятника разница координат $\Delta q \equiv \Delta\theta_{12} = \theta_1 - \theta_2$ и аппроксимируемый лагранжиан примет вид:

$$L = \frac{1}{2}\left(m_1 + m_2\right) l_1^2 \dot{\theta}_1^2 + \frac{1}{2} m_2 l_2^2 \dot{\theta}_2^2 + m_2 l_1 l_2 \dot{\theta}_1 \dot{\theta}_2 \cos\left(\Delta\theta_{12}\right) + V(\Delta\theta_{12})$$

Теорема 1 (Северилов, 2022)

Нётеровская LNN учитывает трансляционную симметрию (закон сохранения импульса). Т.е. лагранжиан системы не изменится, если подвергнуть элементы системы перемещению $\boldsymbol{\xi}$:

$$L(\boldsymbol{\Delta}\tilde{\boldsymbol{q}}, \boldsymbol{\Delta}\dot{\tilde{\boldsymbol{q}}}) = L(\boldsymbol{\Delta}\boldsymbol{q}, \dot{\boldsymbol{q}}), \quad \tilde{\boldsymbol{q}} = \boldsymbol{q} + \boldsymbol{\xi}$$

Теорема 2 (Северилов, 2022)

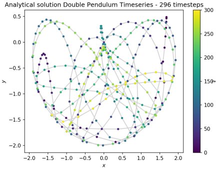
Нётеровская LNN учитывает вращательную симметрию (закон сохранения момента импульса). Т.е. лагранжиан системы не изменится, если подвергнуть элементы системы повороту $\boldsymbol{Q}$:

$$L(\boldsymbol{\Delta}\tilde{\boldsymbol{q}}, \boldsymbol{\Delta}\dot{\tilde{\boldsymbol{q}}}) = L(\boldsymbol{\Delta}\boldsymbol{q}, \dot{\boldsymbol{q}}), \quad \tilde{\boldsymbol{q}} = \boldsymbol{Q}\boldsymbol{q}$$

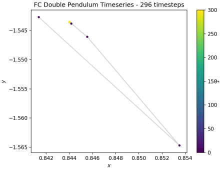
Сравниваемые методы моделирования динамики системы

Аналитическое решение	Метод Рунге-Кутты 4 порядка
Полносвязная нейронная сеть (FC)	$f : \boldsymbol{X} = (\boldsymbol{q}, \dot{\boldsymbol{q}}) \to \boldsymbol{Y}$, $f = \sigma(\ldots \boldsymbol{W}^T \sigma(\boldsymbol{W}^T \boldsymbol{x} + \boldsymbol{b}))$
Нейронная сеть с долгой краткосрочной памятью (LSTM)	$h : \boldsymbol{X} = (\boldsymbol{q}, \dot{\boldsymbol{q}}) \to \boldsymbol{Y}$, $f_t = \sigma_g \left(W_f x_t + U_f h_{t-1} + b_f\right)$ $i_t = \sigma_g \left(W_i x_t + U_i h_{t-1} + b_i\right)$ $o_t = \sigma_g \left(W_o x_t + U_o h_{t-1} + b_o\right)$ $c_t = f_t \circ c_{t-1} + i_t \circ \sigma_c \left(W_c x_t + U_c h_{t-1} + b_c\right)$ $h_t = o_t \circ \sigma_h \left(c_t\right)$
LNN	$f : \boldsymbol{X} = (\boldsymbol{q}, \dot{\boldsymbol{q}}) \to \boldsymbol{L}$, $\ddot{q} = \left(\nabla_{\dot{q}} \nabla_{\dot{q}}^{\top} L\right)^{-1} \left[\nabla_q L - \left(\nabla_q \nabla_{\dot{q}}^{\top} L\right) \dot{q}\right]$
Нётеровская LNN	$f : \boldsymbol{X} = (\boldsymbol{q}, \dot{\boldsymbol{q}}) \to \boldsymbol{V}$, $L = T - V$, $\ddot{q} = \left(\nabla_{\dot{q}} \nabla_{\dot{q}}^{\top} L\right)^{-1} \left[\nabla_q L - \left(\nabla_q \nabla_{\dot{q}}^{\top} L\right) \dot{q}\right]$

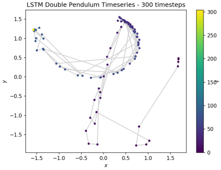
Моделирование динамики системы двойного маятника различными нейронными сетями



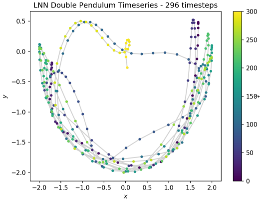
Аналитическое решение



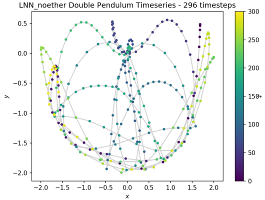
FC



LSTM



LNN



Нётеровская LNN

Добавление априорных знаний о лагранжиане системы повышает точность предсказания динамики системы (траектории движения маятника)

Вычислительный эксперимент: MSE

Выборка состоит из 10 траекторий, сгенерированных на основе решения методом Рунге-Кутты 4 порядка.

FC	1.57 ± 0.53
LSTM	2.42 ± 0.79
LNN	1.32 ± 0.91
Нётеровская LNN	$\mathbf{1.28}$ ± 0.66

Средняя ошибка MSE между предсказанной динамикой системы нейронной сетью и динамикой системы, полученной аналитическим решением.

1. LNN и FC, LSTM: модель, имеющая априорные знания о физике системы имеет выше точность моделирования динамики системы
2. LNN и Нётеровская LNN: добавление дополнительных априорных знаний о физике системы повышает точность моделирования динамики системы

Выносится на защиту

1. Предложена Нётеровская LNN, которая в дополнение к закону сохранения энергии учитывает закон сохранения импульса и момента импульса;
2. Доказаны теоремы, подтверждающие учет закон сохранения импульса и момента импульса;
3. Показано, что добавление априорных знаний о физике системы повышает точность моделирования динамики физической системы.

К публикации

Северилов П.А., Стрижов В.В. Выбор оптимальной модели в задаче моделирования динамики физической системы[2] / 2022

Опубликованные работы

Котлярова Е.В., Северилов П.А., Ивченков Я.П., Мокров П.В., Чеканов М.О., Гасникова Е.В., Шароватова Ю.И. Ускорение работы двухстадийной модели равновесного распределения потоков по сети / Компьютерные исследования и моделирование (том 14), 2022

[2]Код вычислительного эксперимента https://github.com/severilov/master-thesis